

1 **Genetic Association and Transferability for Urinary Albumin-Creatinine Ratio as a**
2 **Marker of Kidney Disease in four Sub-Saharan African Populations and non-continental**
3 **Individuals of African Ancestry**

4
5 Jean-Tristan Brandenburg (PhD)^{1,2}, Wenlong Carl Chen (PhD)^{1,2,3}, Palwende Romuald Boua
6 (PhD)^{1,4}, Melanie Ann Govender (MSc)^{1,5}, Godfred Agongo (PhD)^{6,7}, Lisa K. Micklesfield
7 (PhD)⁸, Hermann Sorgho (PhD)⁴, Stephen Tollman (MD, PhD)⁹, Gershim Asiki (MD,
8 PhD)^{10,11}, Felistas Mashinya (PhD)¹², Scott Hazelhurst (PhD)^{1,13}, Andrew P Morris (PhD)¹⁴,
9 June Fabian (MD)^{15*}, Michèle Ramsay (PhD)^{1,5*} as members of AWI-Gen and the H3Africa
10 Consortium

11
12 ¹ Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University
13 of the Witwatersrand, Johannesburg, South Africa

14 ² Strengthening Oncology Services Research Unit, Faculty of Health Sciences, University of
15 the Witwatersrand, Johannesburg, South Africa

16 ³ National Cancer Registry, National Health Laboratory Service, Johannesburg, South Africa

17 ⁴ Clinical Research Unit of Nanoro, Institut de Recherche en Sciences de la Santé, Burkina
18 Faso

19 ⁵ Division of Human Genetics, National Health Laboratory Service and School of Pathology,
20 Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

21 ⁶ Navrongo Health Research Centre, Navrongo, Ghana

22 ⁷ Department of Biochemistry and Forensic Sciences, School of Chemical and Biochemical
23 Sciences, C. K. Tadam University of Technology and Applied Sciences, Navrongo, Ghana

24 ⁸ SAMRC/Wits Developmental Pathways for Health Research Unit, Faculty of Health
25 Sciences, University of the Witwatersrand, Johannesburg, South Africa

26 ⁹ Medical Research Council/Wits University Rural Public Health and Health Transitions
27 Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of
28 the Witwatersrand, Johannesburg, South Africa

29 ¹⁰ African Population and Health Research Center, Nairobi, Kenya.

30 ¹¹ Department of Women's and Children's Health, Karolinska Institute, Stockholm, Sweden

31 ¹² Department of Pathology and Medical Sciences, School of Health Care Sciences, Faculty of
32 Health Sciences, University of Limpopo, Polokwane, South Africa

33 ¹³ School of Electrical and Information Engineering, University of the Witwatersrand,
34 Johannesburg, South Africa

35 ¹⁴ Centre for Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research,
36 The University of Manchester, Manchester, UK

37 ¹⁵ Wits Donald Gordon Medical Centre, School of Clinical Medicine, Faculty of Health
38 Sciences, University of the Witwatersrand, Johannesburg, South Africa

39
40 *These authors contributed equally to this work and share senior authorship

41
42 Corresponding authors: Dr Jean-Tristan Brandenburg and Prof Michèle Ramsay, Sydney
43 Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the
44 NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Witwatersrand, 9 Jubilee Road, Parktown, Johannesburg, South Africa

45 Tel: +27 11 717 6635

46 E-mail: jean-tristan.brandenburg@wits.ac.za and michele.ramsay@wits.ac.za

47

48 Support: National Institutes of Health (NIH) H3Africa AWI-Gen Collaborative Centre
49 (U54HG006938).

50

51 Running title: GWAS for UACR in African-ancestry populations

52

53 **Keywords (6)**

54 Chronic kidney disease (CKD), genome-wide association study (GWAS), urinary albumin-
55 creatinine ratio (UACR), African diversity, polygenic risk score

56

57 **Abstract (251 words)**

58 **Background:** Genome-wide association studies (GWAS) have predominantly focused on
59 populations of European and Asian ancestry, limiting our understanding of genetic factors
60 influencing kidney disease in Sub-Saharan African (SSA) populations. This study presents the
61 largest GWAS for urinary albumin-to-creatinine ratio (UACR) in SSA individuals, including
62 8,970 participants living in different African regions and an additional 9,705 non-resident
63 individuals of African ancestry from the UK Biobank and African American cohorts.

64 **Methods:** Urine biomarkers and genotype data were obtained from two SSA cohorts (AWI-
65 Gen and ARK), and two non-resident African-ancestry studies (UK Biobank and CKD-Gen
66 Consortium). Association testing and meta-analyses were conducted, with subsequent fine-
67 mapping, conditional analyses, and replication studies. Polygenic scores (PGS) were assessed
68 for transferability across populations.

69 **Results:** Two genome-wide significant ($P < 5 \times 10^{-8}$) UACR-associated loci were identified, one
70 in the *BMP6* region on chromosome 6, in the meta-analysis of resident African individuals,
71 and another in the *HBB* region on chromosome 11 in the meta-analysis of non-resident SSA
72 individuals, as well as the combined meta-analysis of all studies. Replication of previous
73 significant results confirmed associations in known UACR-associated regions, including
74 *THB53*, *GATM*, and *ARL15*. PGS estimated using previous studies from European ancestry,
75 African ancestry, and multi-ancestry cohorts exhibited limited transferability of PGS across
76 populations, with less than 1% of observed variance explained.

77 **Conclusion:** This study contributes novel insights into the genetic architecture of kidney
78 disease in SSA populations, emphasizing the need for conducting genetic research in diverse

79 cohorts. The identified loci provide a foundation for future investigations into the genetic
80 susceptibility to chronic kidney disease in underrepresented African populations. Additionally,
81 there is a need to develop integrated scores using multi-omics data and risk factors specific to
82 the African context to improve the accuracy of predicting disease outcomes.

83

84 **Introduction**

85 Chronic kidney disease (CKD) is a leading risk factor for years of life lost and premature
86 mortality, with a 41.5% relative increase in mortality worldwide from 1990 to 2017 (GBD
87 Chronic Kidney Disease Collaboration et al., 2020; Kovesdy, 2022). The estimated global
88 prevalence of CKD is 9.1% and while predicted to be higher in Sub-Saharan Africa (SSA), the
89 true prevalence and associated risk factors remain understudied (Kaze et al., 2018; GBD
90 Chronic Kidney Disease Collaboration et al., 2020). The Africa Wits-INDEPTH partnership
91 for Genomic Studies (AWI-Gen) cohort, which included ~12,000 participants from four SSA
92 countries in West, East, and Southern Africa, reported overall CKD prevalence as 10.7% (95%
93 confidence interval [CI]: 9.9-11.7), with notable geographic regional differences. The most
94 important risk factors for CKD in SSA were older age, female sex, diabetes, hypertension, and
95 human immunodeficiency virus (HIV) infection (George et al., 2019).

96 Over the past decade, genome-wide association studies (GWAS) have identified numerous
97 genetic loci associated with kidney function disease, namely, estimated glomerular filtration
98 rate [eGFR], serum creatinine, and urine albumin-creatinine ratio [UACR] (Böger et al., 2011;
99 Pattaro et al., 2012, 2016; Teumer et al., 2016, 2019; Hellwege et al., 2019; Tin and Köttgen,
100 2020). The majority of the GWAS for kidney function and disease have examined associations
101 with eGFR, while UACR, as a measure for albuminuria, has been investigated less often
102 (Mahajan et al., 2016; Pattaro et al., 2016; Gorski et al., 2017; Haas et al., 2018; Teumer et al.,
103 2019; Wuttke et al., 2019; Zanetti et al., 2019). A recent GWAS in 564,257 individuals of
104 multi-ancestry origins identified 68 associated risk loci for UACR were identified and proposed
105 priority list of genes to explore as targets for the treatment of albuminuria (Teumer et al., 2019).
106 While the majority of kidney disease-associated risk loci have been identified in studies on
107 participants of European and East Asian ancestry, and the African diaspora (Lee et al., 2018),
108 few have focused on participants living in SSA (Böger et al., 2011; Pattaro et al., 2012; Lin et
109 al., 2019; Morris et al., 2019). Recently, in a study of genetic associations of eGFR in a
110 Ugandan population-based cohort, (Fatumo et al., 2020) replicated the association between
111 eGFR and the *GATM* locus.

112 Replication and transferability of GWAS signals across populations of different ancestries, and
113 specifically with African ancestry populations, tend to be poor despite regional replication
114 often identifying shared associated genomic regions (Pattaro et al., 2012). This may be due to
115 differences in linkage disequilibrium (LD) with the causal variant, allele frequency differences
116 between the populations, underlying population structure, and variabilities in environmental
117 exposures. African populations, with their great genetic diversity and deep evolutionary roots,
118 represent an opportunity for genetic discovery to identify and fine-map disease-associated risk
119 variants (Gomez et al., 2014; Pereira et al., 2021).

120 Polygenic scores (PGS) are used to quantify and stratify populations according to genetic risk.
121 A PGS based on 63 eGFR-associated alleles showed significant association with kidney
122 disease-related phenotypes, such as chronic kidney failure and hypertensive kidney disease in
123 the Million Veteran Study (US) on 192,868 white and non-Hispanic individuals (Hellwege et
124 al., 2019). A PGS based on 64 urine albumin-to-creatinine ratio (UACR) associated alleles was
125 significantly associated with CKD (Teumer et al., 2019). Further analysis revealed positive
126 associations of the PGS with an increased risk of HT and diabetes. However, PGS often
127 translate poorly across different ancestries (Martin et al., 2017; Kamiza et al., 2022; Kachuri
128 et al., 2023). Since most published GWAS for kidney disease and kidney function markers are
129 based on European ancestry populations, the predictive accuracy of models developed from
130 these studies is expected to be significantly diminished for African populations (Adam et al.,
131 2022; Choudhury et al., 2022; Kamiza et al., 2023; Majara et al., 2023).

132 In this study, we present a GWAS for UACR conducted within resident Sub-Saharan African
133 individuals. This population cross-sectional study includes a cohort of 8,970 individuals from
134 four SSA countries from the AWI-Gen study (Ali et al., 2018), the African Research on Kidney
135 Disease (ARK) study (Kalyesubula et al., 2020), with 9,705 individuals of African-ancestry
136 from the UK Biobank (UKB) and African American participants from the CKD-Gen
137 Consortium (Teumer et al., 2019). The primary objectives are to: (1) identify genetic loci
138 associated with UACR as a marker of kidney disease in individuals from SSA and of African
139 ancestry; (2) explore the replication of findings identified in previous GWAS; (3) perform
140 analysis and comparison of PGS derived from non-African and multi-ancestry population
141 studies and evaluate their transferability to African populations.

142

143 **Participants and Methods**

144 **Study participants**

145 *Africa Wits-INDEPTH Partnership for Genomic Research (AWI-Gen)*

146 The study participants are a subset of the population cross-sectional AWI-Gen study (Ramsay
147 et al., 2016; Ali et al., 2018). The study recruited adults primarily between the ages of 40 and
148 80 years from six SSA study sites in West Africa (Nanoro, Burkina Faso and
149 Navrongo, Ghana), East Africa (Nairobi, Kenya) and in South Africa (Bushbuckridge -
150 hereinafter referred to as Agincourt Mpumalanga Province, Dikgale, Limpopo Province; and
151 Soweto, Gauteng Province). All participants were of self-identified black ethnicity. Data
152 collection was described in detail previously (Ali et al., 2018; George et al., 2019). Detailed
153 demographic data, health-related questionnaire data, and anthropometric measurements were
154 collected. Peripheral blood samples and urine samples were collected for biomarker assays (the
155 relevant assays are described below). DNA was extracted from peripheral blood-derived buffy
156 coat samples and used for genotyping. Urine microalbumin was measured using a colorimetric
157 method on the Cobas© 6000/c501 analyzer, and urine creatinine was measured by the modified
158 Jaffe method (Craik et al., 2023). This study was approved by the Human Research Ethics
159 Committee (Medical), University of the Witwatersrand, South Africa (M121029, M170880)
160 and the ethics committees of all participating institutions. All participants provided written
161 informed consent following community engagement and individual consenting processes.

162 *African Research on Kidney Disease (ARK)*

163 The African Research Kidney Disease (ARK) study is a well characterised population-based
164 cohort study of 2021 adults (20–80 years) of self-identified black ethnicity from Agincourt,
165 (Mpumalanga, South Africa) with demographic data, health-related questionnaire data, and
166 anthropometric measurements collected at enrolment (Fabian et al., 2022). Blood and urine
167 were collected for biomarker assays (the relevant assays are described below). DNA was
168 extracted from buffy coat samples and used for genotyping. Urine microalbumin was measured
169 using a colorimetric method on the Cobas© 6000/c501 analyzer, and urine creatinine was
170 measured by the modified Jaffe method (Craik et al., 2023). This study was approved by the
171 Human Research Ethics Committee (Medical), University of the Witwatersrand, South Africa
172 (M160939). All participants provided written informed consent following community

173 engagement and individual consenting processes. The geographical area of recruitment
174 overlaps with the Agincourt sub-cohort of AWI-Gen but there is no overlap in participants.

175 *UK Biobank (UKB)*

176 Individuals of self-reported Caribbean and African ancestry from the UKB were identified for
177 this study. Of this subset of UKB individuals, those with both genotyping and UACR data were
178 retained for the analysis. UACR was derived using urinary levels of albumin and creatinine. In
179 the UKB, albumin was measured using the immuno-turbidimetric analysis method (Randox
180 Biosciences, UK) while creatinine was measured using the enzymatic analysis method
181 (Beckman Coulter, UK) (Casanova et al., 2019).

182 **Phenotype generation and harmonization**

183 UACR was calculated for AWI-Gen and ARK studies using urinary levels of albumin and
184 creatinine as previously described (George et al., 2019; Fabian et al., 2022). Participants with
185 missing values for albumin and creatinine were excluded from this study. We applied filtering
186 criteria similar to those employed by the CKD-Gen consortium (Köttgen and Pattaro, 2020).
187 In cases where the values for urine albumin and urine creatinine fell outside the upper and
188 lower limits of detection, the values were replaced with the respective upper and lower limits:
189 for urine creatinine, the range was 3 to 400 mmol/L and for urine albumin, the range was 3.75
190 to 475 mg/L for AWI-Gen and ARK. For the UKB dataset, the upper limit was 6.7 mg/L for
191 urine albumin. Albuminuria was defined as UACR >3.0mg/mmol.

192 **Genotyping**

193 *AWI-Gen and ARK*

194 Genomic DNA was genotyped using the H3Africa custom genotyping array. The H3Africa
195 custom array was designed as an African-common-variant-enriched GWAS array
196 (<https://www.h3abionet.org/h3africa-chip>) (Illumina) with ~2.3 million single nucleotide
197 polymorphisms (SNPs).

198 *UK-Biobank*

199 Genotyping was performed by Affymetrix on two closely related purpose-designed arrays.
200 ~50,000 participants were genotyped using the UK BiLEVE Axiom array (Resource 149600)
201 and the remaining ~450,000 were genotyped using the UK Biobank Axiom array (Resource
202 149601). The dataset is a combination of results from both arrays. A total of 805,426 markers

203 were released in the genotype data. We extracted individuals with self-reported (Data-Field in
204 dataset 21000) African Ancestry split between African (UKB-African) and Caribbean origins
205 (UKB-Caribbean) from the raw dataset (Casanova et al., 2019).

206 *Quality control*

207 For each dataset, AWI-Gen (Choudhury et al., 2022), ARK and UKB, the following pre-
208 imputation quality control (QC) steps were applied: removal of non-autosomal and
209 mitochondrial SNPs; SNPs with genotype missingness greater than 0.05; minor allele
210 frequency (MAF) less than 0.01; and Hardy-Weinberg equilibrium (HWE) P-value less than
211 0.0001. Individuals were excluded if they had more than 5% overall genotype missingness;
212 heterozygosity lower than 0.150 and higher than 0.343; and discordant genotype and phenotype
213 sex information. We used the GWAS QC workflow of the H3Africa Consortium Pan-African
214 Bioinformatics Network to perform data QC (H3ABioNet H3AGWAS)
215 (<https://github.com/h3abionet/h3agwas>) (Baichoo et al., 2018; Brandenburg et al., 2022a).

216 In our final QC step, we identified and excluded outliers, admixed and related individuals using
217 PCASmart, a feature of the EIGENSOFT software (Price et al., 2006), Admixture software
218 (Alexander et al., 2009) using AGV (Gurdasani et al., 2015) and 1000 Genomes Project data
219 (Auton et al., 2015) and PLINK (Version 1.9) (Purcell et al., 2007; Chang et al., 2015). More
220 detail on filter parameters for each software can be found in **Supplementary Table 1**.

221 *Imputation*

222 Genotype imputation was performed on each dataset separately (AWI-Gen, ARK, and UKB)
223 using the Sanger Imputation Server with the African Genome Resources reference panel
224 (<https://www.sanger.ac.uk/tool/sanger-imputation-service/>). EAGLE2 was used for pre-
225 phasing and the PBWT algorithm was used for imputation (Loh et al., 2016). After imputation,
226 poorly imputed SNPs with info scores less than 0.3 and with a HWE P-value less than 1×10^{-04}
227 were removed. The genomic positions were mapped to GRCh37p11.

228 *Phenotype transformation for association testing*

229 For AWI-Gen, ARK, and UKB datasets, UACR was transformed on the logarithm scale. Linear
230 regression of variables was performed with covariates in R (Version 3.6): $\ln(\text{UACR}) \sim \text{age} +$
231 $\text{sex} + \text{genetic principal components (PCs) 1-5}$. Residuals were extracted and transformed using
232 Rank-Based Inverse Normal Transformation to ensure the normal distribution of residuals
233 (Casanova et al., 2019). PCs were calculated using a sub-set of LD pruned pre-imputed SNPs

234 in PLINK (Version 1.9) (Purcell et al., 2007; Chang et al., 2015). The sub-set was derived by
235 LD pruning using PLINK (Version 1.9) (Purcell et al., 2007; Chang et al., 2015) with an LD
236 (r^2) threshold of 0.2 with windows of 50 kb and 10 kb for step size.

237 *Association testing*

238 Mixed model association testing was performed with imputed genotype probabilities using
239 GEMMA (Version 0.98.1) (Zhou and Stephens, 2012). GEMMA uses a relatedness matrix to
240 account for genetic structure and relatedness between individuals. The relatedness matrix was
241 built with a sub-set of pre-imputed SNPs described above.

242 Mixed model association testing was performed independently on each dataset. A total of nine
243 datasets were tested. The datasets were defined as follows: six datasets for AWI-Gen: AWI-
244 Agincourt, AWI- Dikgale, AWI-Nanoro, AWI-Nairobi, AWI-Navrongo and AWI-Soweto;
245 one dataset for ARK: ARK-Agincourt; and two datasets for UK Biobank: UKB-Caribbean and
246 UKB-African. For each dataset, Quantile-to-quantile plots (QQ-plots) were generated, and
247 inflation factors were calculated using SNPs with MAF>0.01 to verify that the association
248 signals were not inflated due to unaccounted population sub-structure. The genome-wide
249 significance level for novel discovery was considered at $P < 5 \times 10^{-08}$.

250 *CKD-Gen*

251 We used previously published meta-analysis summary statistics from the CKD-Gen
252 Consortium. The CKD-Gen Consortium datasets consist of three meta-analysis summary
253 statistics: 1) CKD-Gen European ancestry individuals (CKD-Gen-EA); 2) CKD-Gen African
254 American ancestry individuals (CKD-Gen-AA); and 3) CKD-Gen Multi-ancestry individuals
255 (CKD-Gen-MA) which include individuals from CKD-Gen-EA and CKD-Gen-AA (Teumer
256 et al., 2019). The CKD-Gen Consortium meta-analysis summary statistics were retrieved from
257 <http://ckdgen.imbi.uni-freiburg.de/>.

258 Briefly, CKD-Gen-AA is a meta-analysis based on 7 studies with African American
259 participants. For each study, genotyping was performed using genome-wide arrays followed
260 by application of study-specific quality filters prior to phasing, imputation, and association
261 analysis software (description can be found in **Supplementary Table 1 and 2** from (Teumer
262 et al., 2019)). Meta-analysis was performed using fixed effects inverse-variance weighted
263 meta-analysis of the study-specific GWAS result files with imputation quality (IQ) score > 0.6
264 and MAC > 10, effective sample size ≥ 100 , and a beta < 10, using METAL (for more details
265 see (Teumer et al., 2019)).

266 **Meta-analysis**

267 Fixed-effect meta-analyses were conducted using the METASOFT software (Han and Eskin,
268 2011). The first meta-analysis (Meta_{SSA}) used the GWAS summary statistics generated from
269 individual-level data from resident SSA populations. This included AWI-Agincourt, AWI-
270 Dikgale, AWI-Nanoro, AWI-Nairobi, AWI-Navrongo, AWI-Soweto and ARK-Agincourt.
271 The second meta-analysis (Meta_{NONRES}) included data from individuals of African ancestry
272 who are not residing in SSA. We used the GWAS summary statistics generated from
273 individual-level data from the UK-Biobank (UKB-African and UKB-Caribbean) and CKD-
274 Gen African American sub-set (CKD-Gen-AA). The third meta-analysis (Meta_{ALL}) consisted
275 of a meta-analysis that pooled the summary statistics of all studies from AWI-Agincourt, AWI-
276 Dikgale, AWI-Nanoro, AWI-Nairobi, AWI-Navrongo, AWI-Soweto, ARK-Agincourt, UKB-
277 African, UKB-Caribbean and CKD-Gen-AA. **Figure 1** outlines the meta-analysis workflow.
278 As a secondary analysis, role of heterogeneity had been investigated between cohorts from
279 different regions of origin by performing separate meta-analyses for residents of Southern
280 African (AWI-Agincourt, AWI-Dikgale, AWI-Soweto, and ARK-Agincourt) and residents of
281 West Africa (AWI-Nanoro, AWI-Navrongo). Random-effects model from METASOFT (Han
282 and Eskin, 2011) take account for potential heterogeneity between study, we performed Meta
283 RE using all dataset (Meta_{ALL}^{RE}) (Borenstein et al., 2010; Adriani Nikolakopoulou et al., 2014).
284 The genome-wide significance level for novel discovery was considered at $P < 5 \times 10^{-08}$.

285

286 **Post association analysis**

287 *Plotting*

288 QQ-plots and Manhattan plots were generated using the FastMan library (Paria et al., 2022)
289 (available at <https://github.com/kaustubhad/fastman>) and the Hudson library (available at
290 <https://github.com/anastasia-lucas/hudson>). These visualizations were created using SNPs with
291 a MAF threshold of 0.01 or more. For regional plots, we utilized the standalone version of the
292 LocusZoom software (Pruim et al., 2010).

293 *Genetic LD reference*

294 For the estimation of the LD reference panel for conditional and joint (COJO) analysis,
295 clumping, and fine-mapping, three LD reference panels were constructed using genotype data
296 from the appropriate datasets. For resident SSA dataset comparisons, the LD reference panel

297 (LD_{SSA}) was constructed using AWI-Gen and ARK individual-level genotype data. For non-
298 resident SSA dataset comparisons, the LD reference panel (LD_{NONRES}) was constructed using
299 UKB individual-level genotype data. For the combined datasets comparison, the LD reference
300 panel (LD_{ALL}) was constructed using AWI-Gen, ARK, and UKB individual-level genotype
301 data.

302 *Fine-mapping and lead SNPs*

303 For each locus with a lead SNP with a P-value below 5×10^{-08} , fine-mapping was conducted
304 using the H3ABioNet H3AGWAS pipeline and implementing a stepwise model selection
305 procedure through GCTA (Yang et al., 2011, 2012; Brandenburg et al., 2022a) to identify
306 independently associated SNPs. Subsequently, we utilized the FINEMAP software (Version
307 1.4) (Benner et al., 2016), considering one causal variant, to define the credible set with 99%
308 confidence using a stochastic approach (Benner et al., 2016).

309 *Conditional analyses (GCTA)*

310 Conditional analyses used the GCTA software implemented within the H3AGWAS pipeline,
311 with summary statistics obtained from the meta-analyses as input. In these analyses, the lead
312 SNPs identified in each meta-analysis were conditioned upon lead SNPs found in previously
313 published studies. Changes in the p-value, both increasing or decreasing significance, of the
314 lead SNP, confirmed a relationship between the two SNPs.

315 *Replication of previous findings*

316 Replication was performed according to the following criteria: 1) Exact replication: if any
317 genome-wide significant lead SNPs found in CKD-Gen-EA and CKD-Gen-MA reached
318 statistical significance ($p < 0.05$) in Meta_{SSA}, Meta_{NONRES} or Meta_{ALL} after Bonferroni correction
319 (A total of 60 independent lead SNPs were identified in the CKD-Gen datasets, of which 55
320 lead SNPs were from CKD-Gen-EA and 57 lead SNPs were from CKD-Gen-MA) and that the
321 lead SNPs have same direction of effect. 2) LD Window replication: for a given genome-wide
322 significant SNP found in the CKD-Gen datasets, SNPs were extracted from Meta_{SSA},
323 Meta_{NONRES} and Meta_{ALL} that are in LD with the said CKD-Gen lead SNP. LD pruning used
324 the clump procedure in PLINK (Version 1.9) ($r^2 = 0.1$, windows size 1000 kb, $P_1 = 5 \times 10^{-08}$, P_2
325 $= 0.1$). The lowest p-value(s) from SNPs within the given LD window were extracted and this
326 LD window was considered statistically significant if the p-value was less than 5×10^{-04} in both
327 datasets. Additionally, the direction of effect between the CKD-Gen and Meta-datasets

328 (Meta_{SSA}, Meta_{NONRES} and Meta_{ALL}) must be consistent. Conditional analyses were performed
329 between the genome-wide significant SNP(s) in CKD-Gen and lead SNP in our meta-analyses
330 to confirm the replication.

331 For replication, the findings from Meta_{SSA} were compared to CKD-Gen-MA and CKD-Gen-
332 EA, and the findings from Meta_{NONRES} and Meta_{ALL} were only compared to CKD-Gen-EA to
333 avoid sample overlaps within the CKD-Gen datasets (as CKD-Gen-AA is contained within
334 CKD-Gen-MA).

335 *Annotation and Expression Quantitative Trait Locus (eQTL) analysis*

336 Functional annotation of genome-wide significant SNPs found in Meta_{SSA}, Meta_{NONRES} and/or
337 Meta_{ALL} was done using the ANNOVAR software (Wang et al., 2010). eQTL analysis was
338 performed using the database of cis-eQTLs in both glomerular and tubulointerstitial tissues,
339 derived from participants in the Nephrotic Syndrome Study Network (NEPTUNE) using SNPs
340 with false discovery rate (FDR) < 0.05 (Han et al. 2023). In this analysis a 1000kb window
341 was defined around each genome-wide significant locus and an eQTL was considered
342 significant if the LD (r^2) was ≥ 0.01 between the lead SNP and significant eQTL, LD
343 computation used the genetics data from the African populations from the 1000 Genomes
344 Project (v5a, hg19) (Auton et al., 2015; Sudmant et al., 2015).

345 *Polygenic Scores*

346 PGS were computed for each dataset independently (AWI-Agincourt, AWI- Dikgale, AWI-
347 Nanoro, AWI-Nairobi, AWI-Navrongo, AWI-Soweto ARK-Agincourt, UKB-African and
348 UKB-Caribbean). The effect sizes from 3 previous studies were used: CKD-Gen-AA, CKD-
349 Gen-MA, and CKD-Gen-EA. PRS-CS (Ge et al., 2019), software that estimates posterior SNP
350 effect sizes by implementing continuous shrinkage (CS) priors, was used to calculate the PGS.
351 As external LD references are required for this analysis, the African LD data derived from the
352 1000 Genomes Project by the PRSs project was used for this purpose (accessible at
353 <https://github.com/getian107/PRSs>). The PGS values were regressed against the residualized
354 UACR value in a linear regression model that adjusted for age, sex, and the first five principal
355 components to assess the performance of PGS.

356

357 **Results**

358 **Study participants and phenotype data**

359 Genomic and phenotypic data were accessible for 7,959 individuals in the AWI-Gen datasets,
360 1,011 individuals in the ARK dataset, and 2,916 individuals in the UK-Biobank dataset with
361 1,205 individuals and 1,711 individuals in UKB-African and UKB-Caribbean respectively
362 (**Supplementary Figure S1**). CKD-Gen AA was a meta-analysis of 7 studies including 6,795
363 individuals in total. Overall, there was a higher prevalence of albuminuria (17.9%; median
364 UACR 1.01mg/mmol) among individuals from the UKB with African and Caribbean ancestry
365 compared to individuals residing in SSA, where notable regional differences were observed.
366 The highest prevalence of albuminuria occurred in AWI-Agincourt, South Africa (14.1%;
367 median UACR 0.59mg/mmol) while the lowest prevalence occurred in AWI-Nanoro, West
368 Africa (prevalence of albuminuria 4.5%, median UACR 0.35 mg/mmol) (**Table 1**).

369 **Meta-analysis**

370 Meta-analyses were conducted to investigate the genetics of UACR in resident Sub-Saharan
371 African datasets ($Meta_{SSA}$) (**Figure 2a**), non-resident Sub-Saharan African datasets
372 ($Meta_{NONRES}$) (**Figure 2b**) and all African ancestry datasets ($Meta_{ALL}$) (**Figure 2c**).

373 No genomic inflation was observed for the individual-dataset association testing performed on
374 the 9 datasets. All genomic inflation factors (λ) were below 1.1. This was visually
375 confirmed on the dataset specific QQ-plots and Manhattan plots (**Supplementary Figure S2**
376 **and Figure S3(a-i)**). Dataset-specific significant findings are reported in **Supplementary**
377 **Table 2 and Supplementary Figure S4(a-c)**.

378 One genome-wide significant locus with the lead SNP rs9505286 ($p=4.3 \cdot 10^{-08}$) was identified
379 in $Meta_{SSA}$ on chromosome 6. (**Figure 3a**). One genome wide significant locus with the lead
380 SNP rs73404549 was identified on chromosome 11 in $Meta_{NONRES}$ ($p=5.6 \cdot 10^{-11}$) and $Meta_{ALL}$
381 ($p=7.7 \cdot 10^{-13}$) (**Table 2, Figure 3b and 3c**)

382 SNP rs9505286 (chr6:7820353) is located in the intronic regions of *BMP6*. Two SNPs were
383 identified in the 95% credible set using FINEMAP (**Figure 3a, Supplementary Table 4**).
384 eQTLs in the region were found to be associated with the expression of two genes *RREB1* and
385 *BMP6* (**Table 2; Supplementary Figure S5; Supplementary Table 3, Supplementary Table**
386 **5**).

387 SNP rs73404549 (chr11:5320654) is located near the *HBE1*, *OR51B4*, and *HBB* genes. This
388 signal is primarily driven by results from West African ancestry datasets in the $Meta_{NONRES}$ and
389 $Meta_{ALL}$ (**Figure 3b and 3c, Supplementary Figure S6**). Notably, this SNP is monomorphic

390 in the Southern African and East African datasets. Furthermore, rs73404549 is in LD with
391 rs334 ($r^2=0.52$; 72,422 bp apart), the SNP that defines the sickle cell mutation (HbS). SNP
392 rs334 was also significant in Meta_{ALL} ($P_{ALL}=8.55 \times 10^{-9}$).

393

394 In the window of 1000kb around rs73404549, SNPs in the region colocalized with gene
395 expression of *TRIM6* and *STIM1* in glomerular and tubulointerstitial tissues (**Table 2,**
396 **Supplementary Table 5**).

397

398 Seven and two SNPs were identified in the 95 % credible set using FINEMAP in Meta_{NONRES}
399 and Meta_{ALL} results respectively (**Supplementary Table 3; Supplementary Figure S6**).

400 **Replication of previous findings**

401 Replication analysis confirmed associations in three previously identified regions in *THBS3*,
402 *SPATA5L1/GATM*, and *ARL15* (**Supplementary Table 3**).

403 In the *THBS3* region, the Meta_{ALL} meta-analysis rs370545 was the lead SNP in our dataset,
404 with a P-value of 1×10^{-04} . However, a conditional analysis using rs2974937 (lead SNP in CKD-
405 Gen-EA) resulted in a decrease in significance level ($P_{\text{conditional_analysis}}=0.85$). This suggests that
406 the association in the *THBS3* region was driven by rs2974937 in Meta_{ALL} even though it was
407 not the lead SNP in this region (**Supplementary Table 4; Supplementary Figure 7**).

408 In the *ARL15* region, a statistically significant association signal was observed in Meta_{ASSA}
409 (rs1664781, $P=1.8 \times 10^{-04}$). Conditional analysis using rs1694068 (lead SNP in CDK-Gen-EA)
410 revealed a reduction in P-value for rs1664781 ($P_{\text{conditional_analysis}}=0.87$), suggesting that
411 rs1694068 and rs1664781 are in LD thus confirming the association in this region
412 (**Supplementary Table 4, Supplementary Figure 8**).

413 In the *SPATA5L1/GATM* region, the Meta_{ALL} meta-analysis identified rs1694067 as the lead
414 SNP in this region with a P-value of 7.0×10^{-05} . Furthermore, the lead SNP rs1153847 identified
415 in CKD-Gen-EA, rs1153847, was present in our dataset, and its association was replicated
416 ($P_{\text{Bonferroni_adjusted}}=0.04$). For the window-based replication, a conditional analysis using
417 rs2467858 (genome-wide significant SNP in CKD-Gen-EA), a reduced P-value was observed
418 ($P_{\text{conditional_analysis}}=0.87$) confirming rs1694067 and rs2467858 are in LD and replicated the
419 CKD-Gen signal. (**Supplementary Table 4; Supplementary Figure 9**).

420 **Polygenic score analyses**

421 The variance explained by the PGS for UACR residuals was between 0% and 0.82%. PGS
422 constructed using the betas from CKD-Gen-EA and CKD-Gen-MA performed better for the

423 non-SSA resident datasets, particularly in the UKB-African, showing the best predictivity (%
424 variance: 0.82, $p=1 \times 10^{-04}$) and statistically significant correlation between the PGS and the
425 UACR residual (**Figure 4, Supplementary Table S6**).

426 Using the PGS constructed from CKD-Gen-MA, ARK-Agincourt (% variance: 0.61, $P=0.01$)
427 and AWI-Agincourt (% variance: 0.58, $P=0.002$) demonstrated better predictivity in SSA
428 populations. PGS constructed from CKD-Gen-AA did not improve the variance explained.
429 Variance explained was lower using PGS constructed from CKD-GEN-AA than CKD-Gen-
430 MA or CKD-Gen-EA.

431

432 **Discussion**

433 This study is the first GWAS for UACR conducted in Sub-Saharan African populations. Two
434 genomic regions were identified to be significantly associated with UACR among 8,970
435 participants from West, East, and Southern Africa and among 9,705 non-resident African-
436 ancestry participants from the UK Biobank and CKD-Gen Consortium.

437 For the first locus, the SNP rs9505286, reached genome-wide significance in resident African
438 individuals (MetaSSA) and is located in the intronic region of *BMP6*. eQTLs in LD with
439 rs9505286 were found to be associated with expression of two genes, namely bone
440 morphogenetic protein 6 (*BMP6*) and ras-responsive element binding protein 1 (*RREB1*). Both
441 genes are plausibly linked with kidney disease. *BMP6* encodes a secreted ligand of the
442 transforming growth factor (TGF-beta) superfamily of proteins, of which TGF-B1 is one of the
443 most important regulators of kidney fibrosis, the pathological hallmark of irreversible loss of
444 kidney function in CKD (Dendooven et al., 2011; Jenkins and Fraser, 2011). TGF-B1 is highly
445 expressed in various fibrotic kidney diseases, including diabetic nephropathy (DN),
446 hypertensive nephropathy, obstructive kidney disease, autosomal dominant polycystic kidney
447 disease, immunoglobulin A nephropathy, crescentic glomerulonephritis, and focal segmental
448 glomerulosclerosis. Because of its pivotal role in mediating kidney fibrosis, TGF-B1 is a
449 potential target for drug discovery, and these results point towards similar potential in African
450 populations for further exploration. *RREB1*, initially identified as a repressor of the
451 angiotensinogen gene, is associated with type 2 diabetes in African Americans with end stage
452 kidney disease (Bonomo et al., 2014). *RREB1* polymorphisms have been shown to interact with
453 *APOL1*, and are implicated in fat distribution and fasting glucose, a potential explanation for
454 the association with type 2 diabetes. As obesity and type 2 diabetes prevalence emerge in many
455 African communities undergoing rapid sociodemographic transition, these findings must

456 inform future work (Bonomo et al., 2014). Unfortunately, neither of the eQTLs has strong
457 linkage disequilibrium (LD) support with the lead SNP rs9505286 (see Supplementary Table
458 5).

459 Variability of kidney function, confounding factors and allele frequency differences between
460 datasets may explain why the rs9505286 signal was not replicated in Meta_{ALL} or Meta_{NONRES}
461 (Marigorta et al., 2018).

462 For the second locus, the SNP rs73404549 was found to be statistically significant in non-
463 resident individuals with African Ancestry (Meta_{NONRES}) and overall (Meta_{ALL}), but not in Sub-
464 Saharan African individuals (Meta_{SSA}). This can be explained by the fact that the variant allele
465 of rs73404549 is extremely rare or absent in East and South African populations. This SNP
466 was found to be in linkage disequilibrium with rs334, the sickle cell trait (HbS) in the *HBB*
467 gene. The HbS mutation has been linked to malaria resistance among heterozygotes, with
468 differences in allele frequency attributed to variations in selection pressures between Bantu-
469 speaking populations in West and South/East Africa (Gurdasani et al., 2019; Choudhury et al.,
470 2020). Notably, sickle cell trait and rs334 had been associated with various kidney function
471 (eGFR) and kidney disease traits, including albuminuria, and chronic and end-stage kidney
472 disease in African, African American and US Hispanic/Latino populations (Naik et al., 2014;
473 Gurdasani et al., 2019; Fatumo et al., 2020; Masimango et al., 2022). Furthermore, an
474 interaction between *APOLI* high-risk genotypes and the sickle cell trait enhances the risk for
475 low eGFR (Masimango et al., 2022).

476 In addition to the *HBB* region, our GWAS revealed transferability of three previously identified
477 signals. Of the 60 UACR-associated loci identified in European and Multi-Ancestry studies,
478 only three were replicated, including variants in *GATM*. This region was also associated with
479 eGFR in a Ugandan population (Fatumo et al., 2020). We also replicated the association with
480 *ARL15* in the region of chromosome 1. *ARL15* is a regulator of Mg²⁺ transport thereby
481 promoting the complex N-glycosylation of cyclin M proteins (CNNM 1-4) and could play a
482 role in the pathogenesis of hypertension mediated via altered tubular handling of magnesium
483 in the kidney (Zolotarov et al., 2021).

484 Allelic heterogeneity is high in African ancestry populations, as demonstrated by the high
485 genetic diversity in our study (**Supplementary Fig. 1**). However, analysis of regional
486 subgroups using meta-analysis (residents of South, West, or East Africa) did not reveal
487 significant population-specific signatures ($p < 5e-8$), likely due to small sample sizes within
488 these subgroups (**Supplementary Fig. 10a, 10b, 3i**). Interestingly, meta-analysis under a
489 random-effects model that allows for heterogeneity in allelic effects between regions

490 (Meta_{ALL}^{RE}) did not improve the detection of specific signals already observed with the fixed-
491 effects methods for HBB (Supplementary Figure S11). Consequently, the heterogeneity
492 observed might be explained primarily by variations in linkage disequilibrium or
493 environmental factors rather than by the effect of a specific allele, such as the presence or
494 absence of sickle cell trait (Adriani Nikolakopoulou et al., 2014; Kuchenbaecker et al., 2019;
495 Choudhury et al., 2020).

496 The transferability of PGS developed using the effect sizes quantified in three previous
497 association studies in European ancestry, African ancestry and multi-ancestry populations
498 showed limited predictability, explaining less than 1% of the variability in UACR. PGS in
499 resident African populations (AWI-Gen and ARK) explained between 0.58% and 0.60% of the
500 variance of UACR compared to UKB-African, where best prediction was observed (0.80%).
501 The poor predictability of UACR using summary statistics derived from African Americans
502 was likely due to the small sample size of the discovery dataset. Unfortunately, there have been
503 few studies on PGS approaches to compare findings with, and the genetic heritability of UACR
504 is relatively low, estimated at 4.3% (Teumer et al., 2019).

505 The limited transferability of PGS and previous GWAS signals across ancestral groups could
506 be due to differences in genetic architecture and/or pleiotropic effects. Different demographic
507 histories and genetic selection pressures between European and African populations could
508 modify the ability to replicate previous GWAS results due to differences in allele frequencies
509 between non-African and African populations, with generally lower LDs in African genomes.
510 Environmental factors and variability in the prevalence and aetiology of kidney - and disease-
511 related risk factors such as diabetes and hypertension (Fatumo et al., 2020) could also influence
512 the genetic architecture of kidney disease in Africans populations (Limou et al., 2014; Teumer
513 et al., 2019; Brandenburg et al., 2022b). Selection pressures have increased the frequencies
514 *APOLI* kidney risk variants and HbS due to their protective properties in areas of Africa where
515 trypanosomiasis and malaria are endemic. This may have contributed to shaping genetic
516 susceptibility to kidney disease in African individuals. In our study, the *APOLI* gene region
517 did not exhibit significant associations with UACR. The indel rs71785313 was not imputed
518 using the African Sanger reference for imputation, and a specific study had previously been
519 published to describe *APOLI* variant distribution in the AWI-Gen dataset using other
520 imputation panels, but the locus did not reach genomic significance ($5e^{-8}$) for association with
521 eGFR and UACR (Brandenburg et al., 2022b).

522 While the burden of CKD in SSA is high, it is noteworthy that no prior GWAS on UACR has
523 been conducted on the continent. Despite its uniqueness, our study is limited by its relatively

524 modest sample size, which impacts statistical power to detect small-effect associations
525 reaching genome-wide significance thresholds. Kidney and disease markers were measured at
526 a single time point, and spot urine albumin and creatine levels are sensitive to incident
527 infections and other environmental factors that could affect the prevalence of albuminuria.

528 It is important to note that our study populations are mainly treatment naïve in relation to
529 kidney disease and other cardiometabolic conditions, which may be an advantage in detecting
530 genetic associations (Pereira et al., 2021). Other studies, based on lipid-associated loci,
531 attributed non-transferability of associated loci to pleiotropic effects, gene-environment
532 interactions, and also to variability in allele frequencies and LD patterns (Kuchenbaecker et al.,
533 2019; Choudhury et al., 2022), as we hypothesize for UACR.

534 In conclusion, this study describes genetic associations with UACR in a unique SSA cohort
535 and non-resident individuals with African ancestry. CKD in African populations remains
536 understudied but from available data, hypertension, rather than diabetes is the most commonly
537 associated risk factor and in some regions, up to 60% of people with CKD do not have an
538 associated “traditional” risk factor common to high-income settings, suggesting alternate
539 underlying molecular pathways or aetiologies for CKD (Kalyesubula et al., 2018; Nakanga et
540 al., 2019; Muiru et al., 2020). Our study identified two novel SNPs associated with UACR in
541 populations of African ancestry. We further replicated three known UACR-associated loci.
542 Regional genetic diversity due to different selection pressures appear to play a role in the
543 genetic aetiology of CKD across the African continent. These factors likely contribute to the
544 limited transferability of previous association signals and the poor transfer of polygenic scores
545 developed in non-African populations to African populations. Larger genomic studies are
546 necessary to better understand the genetic architecture of kidney function and chronic kidney
547 disease across different African populations and inform region-specific kidney risk profiles.
548 As demonstrated in this study, the low genetic heritability of UACR limits the predictive power
549 of genetic risk scores (GRS) for kidney disease in our setting. It is critical for future research
550 to address these gaps by modelling integrative risk scores that incorporate locally relevant
551 clinical risk factors that are powerful predictors of kidney disease, multiple kidney phenotypes
552 ($eGFR_{\text{cystatin C}}$, $eGFR_{\text{creatinine}}$, $eGFR_{\text{creatinine} + \text{cystatin C}}$, albuminuria, Blood urea nitrogen), using
553 multi-omics (Eddy et al., 2020), and the impacts of African-specific genetic risk for kidney
554 disease, such as *APOLI* high-risk genotypes and sickle cell trait or disease (Naik et al., 2014;
555 Friedman and Pollak, 2016; Brandenburg et al., 2022b).

556

557 **Disclosure statement**

558 The authors have nothing to disclose.

559

560 **Ethics Statement**

561 This study received approval from the Human Research Ethics Committee (Medical),
562 University of the Witwatersrand/South Africa (M121029; M170880; M160939), the approval
563 of the Centre Muraz Institutional Ethics Committee/Burkina Faso (015-2014/CE- CM) and the
564 approval of the National Ethics Committee For Health Research/Burkina Faso (2014-08-096),
565 the Ghana Health Service Ethics Review Committee (ID No: GHS-ERC:05/05/2015), the
566 AMREF Health Ethics and Scientific Review Committee in Kenya (approval no. P114/2014)
567 and the Navrongo Institutional Review Board (ID No: NHRCIRB178). All the participants
568 signed an Informed Consent Form before any study procedure was performed.

569

570 **Conflict of Interest**

571 None

572

573 **Acknowledgement**

574 We acknowledge the sterling contributions of our field workers, phlebotomists, laboratory
575 scientists, administrators, data personnel, and all other staff who contributed to the AWI-Gen
576 and ARK data and sample collection, processing, storage, and shipping, and the participants
577 who generously agreed to be part of the studies. The UK Biobank kindly shared data under the
578 UK Biobank Resource Application Number 63215 and approved submission of the paper.

579

580

581 **Funding**

582 This study was funded by the National Institutes of Health (NIH) through the H3Africa AWI-
583 Gen project (NIH grant number U54HG006938)). AWI-Gen is supported by the National
584 Human Genome Research Institute (NHGRI), Eunice Kennedy Shriver National Institute of
585 Child Health & Human Development (NICHD), and Office of the Director (OD) at the
586 National Institutes of Health. PB is funded by the National Research Foundation/The World
587 Academy of Sciences “African Renaissance Doctoral Fellowship” (Grant no. 100004). JTB is
588 supported by grants from the National Human Genome Research Institute (U54HG006938) as
589 part of the H3A Consortium (AWI-Gen); and the Science For African Foundation - REACCT-
590 CAN Grant (Del-22-008). MR is the South African Research Chair in Genomics and
591 Bioinformatics of African populations hosted by the University of the Witwatersrand
592 (SARChI), funded by the South African Department of Science and Innovation, and

593 administered by the National Research Foundation. WCC is supported by the Research
594 Networks for Health Innovations in sub-Saharan Africa Funding Initiative of the German
595 Federal Ministry of Education and Research (RHISSA) Grant and the Science For African
596 Foundation - REACCT-CAN Grant (Del-22-008). The ARK study was jointly funded by the
597 South African MRC, MRC UK (via the Newton Fund), and GSK Africa Non-Communicable
598 Disease Open Lab (via a supporting grant; project number 074). The views expressed in this
599 manuscript do not necessarily reflect the views of the funders.

600 **Author Contributions**

601 JTB and MR designed the study, JTB performed the analyses and wrote the first draft. WCC,
602 MR, AM, SH and JF made significant contributions to guiding the analyses and interpreting
603 the results, as well as writing and reviewing the manuscript. All authors read, edited and
604 approved the manuscript.

605

606 **Data availability statement**

607 Summary statistics are available in the GWAS catalog (XXX) and scripts are available from
608 the corresponding authors. The AWI-Gen data set is available from the European Genome-
609 phenome Archive (EGA) database (<https://ega-archive.org/>), accession number
610 EGAS00001002482 (phenotype dataset: EGAD00001006425; genotype dataset:
611 EGAD00010001996). The availability of these datasets is subject to controlled access through
612 the Data and Biospecimen Access Committee of the H3Africa Consortium. ARK data is
613 available in WIREDSpace repository on request (<https://wiredspace.wits.ac.za/>),
614 <https://doi.org/10.54223/uniwitwatersrand-10539-33016> 35

615 Data are available under the terms of the Creative Commons Zero "No rights reserved" data
616 waiver (CC0 1.0 Public domain dedication). The processed data generated in this study are
617 provided in Supplementary Material. Permission was obtained to access the genotype and
618 phenotype dataset for UKBB (research project number: 63215).

619 Publicly available databases include Teumer et al. in CKD-GEN consortium website:
620 <https://ckdgen.imbi.uni-freiburg.de/>.

621

622 **Tables**

623

624 **Table 1: Study participants and phenotype data.** Participant characteristics for each AWI-
625 Gen study site, ARK-Agincourt and UKB-African and UKB-Caribbean, with phenotype
626 distributions of UACR (median) and covariables used in the study.

627

Dataset	Sample size	Country of residence	¹ Age (years)	Males (%)	² UACR	³ Albuminuria (%)
Non-resident in ⁴SSA	9,705					
⁵ CKD-Gen-AA	6,795	USA				
⁶ UKB-African	1,205	UK	51.8	48.5	1.10 (1.70)	20.0
⁷ UKB-Caribbean	1,711	UK	52.8	59.6	0.93 (1.44)	16.4
⁸ UKB-All	2,916	UK	52.4	55.0	1.01 (1.55)	17.9
Resident in SSA	8,970					
⁹ AWI-Nanoro	1,702	Burkina Faso	49.6	50.8	0.35 (0.44)	4.9
¹⁰ AWI-Navrongo	1,548	Ghana	51.1	54.4	0.41(0.48)	6.7
¹¹ AWI-Nairobi	1,481	Kenya	48.8	54.0	0.63 (0.67)	11.4
¹² AWI-Agincourt	1,545	South Africa	54.3	59.8	0.59 (0.99)	14.1
¹³ AWI-Dikgale	917	South Africa	52.1	68.3	0.63 (0.51)	10.4
¹⁴ AWI-Soweto	766	South Africa	49.5	100.0	0.36 (0.57)	11.4
AWI-All	7,959		50.9	51.0	0.48 (0.51)	9.5
¹⁵ ARK-Agincourt	1,011	South Africa	38.8	58.0	0.57 (0.94)	13.4
Total	17,664		50.3	52.6	0.59 (0.85)	11.9

628

¹data reported as the mean.

629

²UACR: urine albumin: creatinine ratio: mg/mmol; reported as median (interquartile range).

630

³Albuminuria: UACR>3.0mg/mmol.

631

⁴SSA: Sub-Saharan Africa.

632

⁵CKD-Gen Consortium: African American Ancestry individuals. CKD-GEN-AA data are summary statistics of meta-analysis from Teumer et al. downloadable at the CKD-GEN consortium website (<http://ckdgen.imbi.uni-freiburg.de/>), information relative to samples had been reported in relevant papers.

633

⁶UKB Biobank: Individuals of self-reported African Ancestry; ⁷self-reported Caribbean Ancestry; ⁸self-reported African and Caribbean Ancestry.

634

⁹Africa Wits -INDEPTH partnership for Genomic Studies (AWI): individuals of African Ancestry from ^{9,10}West Africa; ¹¹East Africa; ¹²-

635

¹⁴South Africa.

636

¹⁵African Research on Kidney Disease (ARK) Study: South Africa

637

638

639

640 **Table 2:** Lead genome-wide significantly associated SNPs for sub-Saharan African population meta-analysis (MetASSA), non-resident African
 641 ancestry population meta-analysis (Meta_{NONRES}) and the combined African ancestry population meta-analysis (Meta_{ALL}).
 642

Chr	bp	rsID	Genes	eQTL (genes)	EA	NEA	MetASSA			Meta _{NONRES}			Meta _{ALL}		
							beta (se)	P	EAF	beta (se)	P	EAF	beta (se)	P (beta, se)	EAF
6	7820353	rs9505286	intronic: BMP6	<i>RREB1, BM P6</i>	T	C	0.19 (0.03)	4.3e- 08	0.95	-0.05 (0.07)	0.17	0.96	-0.13 (0.03)	7.2e-07	0.96
11	5320654	rs73404549	intergenic: HBE1, OR51B4	<i>TRIM6, TRI M5, STIM1, OR51E2.A C104389.2 8, PRKCDB P, UBQLNL , OR51E1.H BE1, AC087 380.14, OR5 2K3P, RP11 - 23F23.3, TR IM6- TRIM34, TR IM34, RP11 -290F24.6</i>	T	C	0.17 (0.06)	3.2e- 03	0.05	0.21 (0.06)	5.6e-11	0.05	0.2 (0.04)	7.7e-13	0.05

643 Chr: chromosome (hg19)
 644 bp: position coordinates
 645 rsID: SNPs identification
 646 Genes: genes annotation using ANNOVAR
 647 eQTL: gene's transcript is significantly associated with nearby SNP positions (1000kb and LD>0.01)
 648 EA: Effect allele
 649 NEA : non effect allele
 650 beta (se): effect size and standard error
 651 P: significance of association
 652 EAF: Effect allele frequency

653
654
655

656 **Figure Legends**

657

658 **Figure 1:** Study design showing data sources, the analysis strategy and post-GWAS analysis
659 approach.

660

661 **Figure 2:** Manhattan plot - GWAS of UACR in the (a) Meta_{ASSA} (b) Meta_{NONRES} (c) Meta_{ALL}
662 datasets using the fixed effect model. Lead genome-wide significant SNPs ($P < 5 \times 10^{-08}$) and
663 gene annotations are highlighted.

664

665 **Figure 3:** Regional plot using LocusZoom of genome-wide significant SNPs found in meta-
666 analyses using the fixed effect model, (a) rs9505286 from the result of Meta_{ASSA}, (b) rs9966824
667 from the result Meta_{NONRES} (c) rs9966824 from the result Meta_{ALL}.

668

669 **Figure 4.** Percent variance explained between PGS and residual phenotypes computed using
670 age, sex and 5 PCs. Key: The negative relationship between PGS and the phenotype in the
671 result of the linear model, * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$. Details in Supplementary
672 Table 6

673

674

675 **References**

676 Adam, Y., Sadeeq, S., Kumuthini, J., Ajayi, O., Wells, G., Solomon, R., et al. (2022). Polygenic
677 Risk Score in African populations: progress and challenges. *F1000Res* 11, 175. doi:
678 10.12688/f1000research.76218.2

679 Adriani Nikolakopoulou, Dimitris Mavridis, and Georgia Salanti (2014). How to interpret
680 meta-analysis models: fixed effect and random effects meta-analyses. *Evid Based Ment*
681 *Health* 17, 64. doi: 10.1136/eb-2014-101794

682 Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of
683 ancestry in unrelated individuals. *Genome Res* 19, 1655–1664. doi:
684 10.1101/gr.094052.109

685 Ali, S. A., Soo, C., Agongo, G., Alberts, M., Amenga-Etego, L., Boua, R. P., et al. (2018).
686 Genomic and environmental risk factors for cardiometabolic diseases in Africa:
687 methods used for Phase 1 of the AWI-Gen population cross-sectional study. *Global*
688 *Health Action* 11, 1507133. doi: 10.1080/16549716.2018.1507133

- 689 Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R.,
690 et al. (2015). A global reference for human genetic variation. *Nature* 526, 68–74. doi:
691 10.1038/nature15393
- 692 Baichoo, S., Souilmi, Y., Panji, S., Botha, G., Meintjes, A., Hazelhurst, S., et al. (2018).
693 Developing reproducible bioinformatics analysis workflows for heterogeneous
694 computing environments to support African genomics. *BMC Bioinformatics* 19, 457.
695 doi: 10.1186/s12859-018-2446-1
- 696 Benner, C., Spencer, C. C. A., Havulinna, A. S., Salomaa, V., Ripatti, S., and Pirinen, M.
697 (2016). FINEMAP: efficient variable selection using summary data from genome-wide
698 association studies. *Bioinformatics (Oxford, England)* 32, 1493–1501. doi:
699 10.1093/bioinformatics/btw018
- 700 Böger, C. A., Chen, M.-H., Tin, A., Olden, M., Köttgen, A., de Boer, I. H., et al. (2011). CUBN
701 is a gene locus for albuminuria. *Journal of the American Society of Nephrology : JASN*
702 22, 555–570. doi: 10.1681/ASN.2010060598
- 703 Bonomo, J. A., Guan, M., Ng, M. C. Y., Palmer, N. D., Hicks, P. J., Keaton, J. M., et al. (2014).
704 The ras responsive transcription factor RREB1 is a novel candidate gene for type 2
705 diabetes associated end-stage kidney disease. *Hum Mol Genet* 23, 6441–6447. doi:
706 10.1093/hmg/ddu362
- 707 Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2010). A basic
708 introduction to fixed-effect and random-effects models for meta-analysis. *Research*
709 *Synthesis Methods* 1, 97–111. doi: 10.1002/jrsm.12
- 710 Brandenburg, J.-T., Clark, L., Botha, G., Panji, S., Baichoo, S., Fields, C., et al. (2022a).
711 H3AGWAS: a portable workflow for genome wide association studies. *BMC*
712 *Bioinformatics* 23, 498. doi: 10.1186/s12859-022-05034-w
- 713 Brandenburg, J.-T., Govender, M. A., Winkler, C. A., Boua, P. R., Agongo, G., Fabian, J., et
714 al. (2022b). Apolipoprotein L1 High-Risk Genotypes and Albuminuria in Sub-Saharan
715 African Populations. *Clin J Am Soc Nephrol* 17, 798–808. doi: 10.2215/CJN.14321121
- 716 Casanova, F., Tyrrell, J., Beaumont, R. N., Ji, Y., Jones, S. E., Hattersley, A. T., et al. (2019).
717 A genome-wide association study implicates multiple mechanisms influencing raised

- 718 urinary albumin–creatinine ratio. *Human Molecular Genetics* 28, 4197–4207. doi:
719 10.1093/hmg/ddz243
- 720 Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015).
721 Second-generation PLINK: rising to the challenge of larger and richer datasets.
722 *Gigascience* 4, 7. doi: 10.1186/s13742-015-0047-8
- 723 Choudhury, A., Aron, S., Botigué, L. R., Sengupta, D., Botha, G., Bensellak, T., et al. (2020).
724 High-depth African genomes inform human migration and health. *Nature* 586, 741–
725 748. doi: 10.1038/s41586-020-2859-7
- 726 Choudhury, A., Brandenburg, J.-T., Chikowore, T., Sengupta, D., Boua, P. R., Crowther, N.
727 J., et al. (2022). Meta-analysis of sub-Saharan African studies provides insights into
728 genetic architecture of lipid traits. *Nature Communications* 13, 2578. doi:
729 10.1038/s41467-022-30098-w
- 730 Craik, A., Gondwe, M., Mayindi, N., Chipungu, S., Khoza, B., Gómez-Olivé, X., et al.
731 (2023). Forgotten but not gone in rural South Africa: Urinary schistosomiasis and
732 implications for chronic kidney disease screening in endemic countries [version 3; peer
733 review: 2 approved]. *Wellcome Open Research* 8. doi:
734 10.12688/wellcomeopenres.18650.3
- 735 Dendooven, A., van Oostrom, O., van der Giezen, D. M., Leeuwis, J. W., Snijckers, C., Joles,
736 J. A., et al. (2011). Loss of endogenous bone morphogenetic protein-6 aggravates renal
737 fibrosis. *Am J Pathol* 178, 1069–1079. doi: 10.1016/j.ajpath.2010.12.005
- 738 Eddy, S., Mariani, L. H., and Kretzler, M. (2020). Integrated multi-omics approaches to
739 improve classification of chronic kidney disease. *Nature Reviews Nephrology* 16, 657–
740 668. doi: 10.1038/s41581-020-0286-5
- 741 Fabian, J., Gondwe, M., Mayindi, N., Chipungu, S., Khoza, B., Gaylard, P., et al. (2022).
742 Chronic kidney disease (CKD) and associated risk in rural South Africa: a population-
743 based cohort study [version 2; peer review: 2 approved]. *Wellcome Open Research* 7.
744 doi: 10.12688/wellcomeopenres.18016.2

- 745 Fatumo, S., Chikowore, T., Kalyesubula, R., Nsubuga, R., Asiki, G., Nashiru, O., et al. (2020).
746 Discovery and fine-mapping of kidney function loci in first genome-wide association
747 study in Africans. doi: 10.1101/2020.06.09.142463
- 748 Friedman, D. J., and Pollak, M. R. (2016). Apolipoprotein L1 and Kidney Disease in African
749 Americans. *Trends in endocrinology and metabolism: TEM* 27, 204–215. doi:
750 10.1016/j.tem.2016.02.002
- 751 GBD Chronic Kidney Disease Collaboration, B., Purcell, C. A., Levey, A. S., Smith, M.,
752 Abdoli, A., Abebe, M., et al. (2020). Global, regional, and national burden of chronic
753 kidney disease, 1990–2017: a systematic analysis for the Global Burden of Disease
754 Study 2017. *Lancet (London, England)* 395, 709–733. doi: 10.1016/S0140-
755 6736(20)30045-3
- 756 Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A., and Smoller, J. W. (2019). Polygenic prediction
757 via Bayesian regression and continuous shrinkage priors. *Nature Communications* 10,
758 1776. doi: 10.1038/s41467-019-09718-5
- 759 George, J. A., Brandenburg, J.-T., Fabian, J., Crowther, N. J., Agongo, G., Alberts, M., et al.
760 (2019). Kidney damage and associated risk factors in rural and urban sub-Saharan
761 Africa (AWI-Gen): a cross-sectional population study. *The Lancet Global Health* 7,
762 e1632–e1643. doi: 10.1016/S2214-109X(19)30443-7
- 763 Gomez, F., Hirbo, J., and Tishkoff, S. A. (2014). Genetic variation and adaptation in Africa:
764 implications for human evolution and disease. *Cold Spring Harbor perspectives in*
765 *biology* 6, a008524–a008524. doi: 10.1101/cshperspect.a008524
- 766 Gorski, M., van der Most, P. J., Teumer, A., Chu, A. Y., Li, M., Mijatovic, V., et al. (2017).
767 1000 Genomes-based meta-analysis identifies 10 novel loci for kidney function.
768 *Scientific reports* 7, 45040. doi: 10.1038/srep45040
- 769 Gurdasani, D., Carstensen, T., Fatumo, S., Chen, G., Franklin, C. S., Prado-Martinez, J., et al.
770 (2019). Uganda Genome Resource Enables Insights into Population History and
771 Genomic Discovery in Africa. *Cell* 179, 984–1002.e36. doi: 10.1016/j.cell.2019.10.004

- 772 Gurdasani, D., Carstensen, T., Tekola-Ayele, F., Pagani, L., Tachmazidou, I., Hatzikotoulas,
773 K., et al. (2015). The African Genome Variation Project shapes medical genetics in
774 Africa. *Nature* 517, 327–332. doi: 10.1038/nature13997
- 775 Haas, M. E., Aragam, K. G., Emdin, C. A., Bick, A. G., Hemani, G., Davey Smith, G., et al.
776 (2018). Genetic Association of Albuminuria with Cardiometabolic Disease and Blood
777 Pressure. *American Journal of Human Genetics* 103, 461–473. doi:
778 10.1016/j.ajhg.2018.08.004
- 779 Han, B., and Eskin, E. (2011). Random-effects model aimed at discovering associations in
780 meta-analysis of genome-wide association studies. *American journal of human genetics*
781 88, 586–598. doi: 10.1016/j.ajhg.2011.04.014
- 782 Han, S. K., McNulty, M. T., Benway, C. J., Wen, P., Greenberg, A., Onuchic-Whitford, A. C.,
783 et al. (2023). Mapping genomic regulation of kidney disease and traits through high-
784 resolution and interpretable eQTLs. *Nature Communications* 14, 2229. doi:
785 10.1038/s41467-023-37691-7
- 786 Hellwege, J. N., Velez Edwards, D. R., Giri, A., Qiu, C., Park, J., Torstenson, E. S., et al.
787 (2019). Mapping eGFR loci to the renal transcriptome and phenome in the VA Million
788 Veteran Program. *Nature Communications* 10, 3842. doi: 10.1038/s41467-019-11704-
789 w
- 790 Jenkins, R. H., and Fraser, D. J. (2011). BMP-6 emerges as a potential major regulator of
791 fibrosis in the kidney. *Am J Pathol* 178, 964–965. doi: 10.1016/j.ajpath.2010.12.010
- 792 Kachuri, L., Chatterjee, N., Hirbo, J., Schaid, D. J., Martin, I., Kullo, I. J., et al. (2023).
793 Principles and methods for transferring polygenic risk scores across global populations.
794 *Nature Reviews Genetics*. doi: 10.1038/s41576-023-00637-2
- 795 Kalyesubula, R., Fabian, J., Nakanga, W., Newton, R., Ssebunnya, B., Prynne, J., et al. (2020).
796 How to estimate glomerular filtration rate in sub-Saharan Africa: design and methods
797 of the African Research into Kidney Diseases (ARK) study. *BMC Nephrology* 21, 20.
798 doi: 10.1186/s12882-020-1688-0

- 799 Kalyesubula, R., Hau, J. P., Asiki, G., Ssebunya, B., Kusemererwa, S., Seeley, J., et al. (2018).
800 Impaired renal function in a rural Ugandan population cohort. *Wellcome Open Res* 3,
801 149. doi: 10.12688/wellcomeopenres.14863.3
- 802 Kamiza, A. B., Toure, S. M., Vujkovic, M., Machipisa, T., Soremekun, O. S., Kintu, C., et al.
803 (2022). Transferability of genetic risk scores in African populations. *Nature Medicine*
804 28, 1163–1166. doi: 10.1038/s41591-022-01835-x
- 805 Kamiza, A. B., Touré, S. M., Zhou, F., Soremekun, O., Cissé, C., Wélé, M., et al. (2023). Multi-
806 trait discovery and fine-mapping of lipid loci in 125,000 individuals of African
807 ancestry. *Nature Communications* 14, 5403. doi: 10.1038/s41467-023-41271-0
- 808 Kaze, A. D., Ilori, T., Jaar, B. G., and Echouffo-Tcheugui, J. B. (2018). Burden of chronic
809 kidney disease on the African continent: a systematic review and meta-analysis. *BMC*
810 *Nephrology* 19, 125. doi: 10.1186/s12882-018-0930-5
- 811 Köttgen, A., and Pattaro, C. (2020). The CKDGen Consortium: ten years of insights into the
812 genetic basis of kidney function. *Kidney International* 97, 236–242. doi:
813 10.1016/j.kint.2019.10.027
- 814 Kovesdy, C. P. (2022). Epidemiology of chronic kidney disease: an update 2022. *Kidney Int*
815 *Suppl (2011)* 12, 7–11. doi: 10.1016/j.kisu.2021.11.003
- 816 Kuchenbaecker, K., Telkar, N., Reiker, T., Walters, R. G., Lin, K., Eriksson, A., et al. (2019).
817 The transferability of lipid loci across African, Asian and European cohorts. *Nature*
818 *Communications* 10. doi: 10.1038/s41467-019-12026-7
- 819 Lee, J., Lee, Y., Park, B., Won, S., Han, J. S., and Heo, N. J. (2018). Genome-wide association
820 analysis identifies multiple loci associated with kidney disease-related traits in Korean
821 populations. *PLOS ONE* 13, e0194044.
- 822 Limou, S., Nelson, G. W., Kopp, J. B., and Winkler, C. A. (2014). APOL1 kidney risk alleles:
823 population genetics and disease associations. *Advances in chronic kidney disease* 21,
824 426–433. doi: 10.1053/j.ackd.2014.06.005

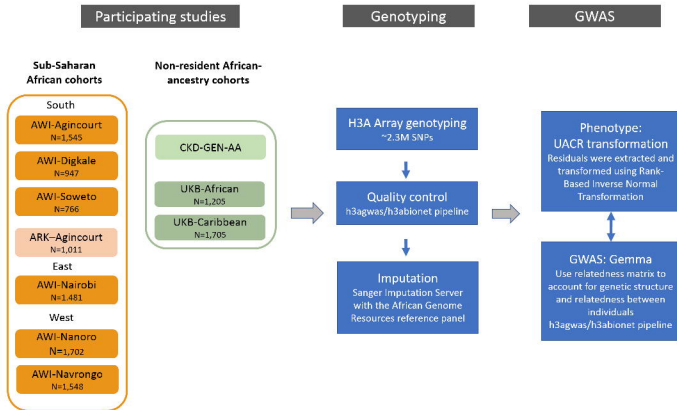
- 825 Lin, B. M., Nadkarni, G. N., Tao, R., Graff, M., Fornage, M., Buyske, S., et al. (2019). Genetics
826 of Chronic Kidney Disease Stages Across Ancestries: The PAGE Study. *Frontiers in*
827 *genetics* 10, 494. doi: 10.3389/fgene.2019.00494
- 828 Loh, P.-R., Danecek, P., Palamara, P. F., Fuchsberger, C., A Reshef, Y., K Finucane, H., et al.
829 (2016). Reference-based phasing using the Haplotype Reference Consortium panel.
830 *Nature Genetics* 48, 1443–1448. doi: 10.1038/ng.3679
- 831 Mahajan, A., Rodan, A. R., Le, T. H., Gaulton, K. J., Haessler, J., Stilp, A. M., et al. (2016).
832 Trans-ethnic Fine Mapping Highlights Kidney-Function Genes Linked to Salt
833 Sensitivity. *American journal of human genetics* 99, 636–646. doi:
834 10.1016/j.ajhg.2016.07.012
- 835 Majara, L., Kalungi, A., Koen, N., Tsuo, K., Wang, Y., Gupta, R., et al. (2023). Low and
836 differential polygenic score generalizability among African populations due largely to
837 genetic diversity. *Human Genetics and Genomics Advances* 4, 100184. doi:
838 10.1016/j.xhgg.2023.100184
- 839 Marigorta, U. M., Rodríguez, J. A., Gibson, G., and Navarro, A. (2018). Replicability and
840 Prediction: Lessons and Challenges from GWAS. *Trends Genet* 34, 504–517. doi:
841 10.1016/j.tig.2018.03.005
- 842 Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., et al.
843 (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse
844 Populations. *Am J Hum Genet* 100, 635–649. doi: 10.1016/j.ajhg.2017.03.004
- 845 Masimango, M. I., Jadoul, M., Binns-Roemer, E. A., David, V. A., Sumaili, E. K., Winkler, C.
846 A., et al. (2022). APOL1 Renal Risk Variants and Sickle Cell Trait Associations With
847 Reduced Kidney Function in a Large Congolese Population-Based Study. *Kidney Int*
848 *Rep* 7, 474–482. doi: 10.1016/j.ekir.2021.09.018
- 849 Morris, A. P., Le, T. H., Wu, H., Akbarov, A., van der Most, P. J., Hemani, G., et al. (2019).
850 Trans-ethnic kidney function association study reveals putative causal genes and effects
851 on kidney-specific disease aetiologies. *Nature communications* 10, 29. doi:
852 10.1038/s41467-018-07867-7

- 853 Muiru, A. N., Charlebois, E. D., Balzer, L. B., Kwarisiima, D., Elly, A., Black, D., et al. (2020).
854 The epidemiology of chronic kidney disease (CKD) in rural East Africa: A population-
855 based study. *PLoS One* 15, e0229649. doi: 10.1371/journal.pone.0229649
- 856 Naik, R. P., Derebail, V. K., Grams, M. E., Franceschini, N., Auer, P. L., Peloso, G. M., et al.
857 (2014). Association of Sickle Cell Trait With Chronic Kidney Disease and Albuminuria
858 in African Americans. *JAMA* 312, 2115. doi: 10.1001/jama.2014.15063
- 859 Nakanga, W. P., Prynne, J. E., Banda, L., Kalyesubula, R., Tomlinson, L. A., Nyirenda, M., et
860 al. (2019). Prevalence of impaired renal function among rural and urban populations:
861 findings of a cross-sectional study in Malawi. *Wellcome Open Res* 4, 92. doi:
862 10.12688/wellcomeopenres.15255.1
- 863 Paria, S. S., Rahman, S. R., and Adhikari, K. (2022). fastman: A fast algorithm for visualizing
864 GWAS results using Manhattan and Q-Q plots. *bioRxiv*. doi:
865 10.1101/2022.04.19.488738
- 866 Pattaro, C., Köttgen, A., Teumer, A., Garnaas, M., Böger, C. A., Fuchsberger, C., et al. (2012).
867 Genome-wide association and functional follow-up reveals new loci for kidney
868 function. *PLoS genetics* 8, e1002584–e1002584. doi: 10.1371/journal.pgen.1002584
- 869 Pattaro, C., Teumer, A., Gorski, M., Chu, A. Y., Li, M., Mijatovic, V., et al. (2016). Genetic
870 associations at 53 loci highlight cell types and biological pathways relevant for kidney
871 function. *Nature communications* 7, 10023. doi: 10.1038/ncomms10023
- 872 Pereira, L., Mutesa, L., Tindana, P., and Ramsay, M. (2021). African genetic diversity and
873 adaptation inform a precision medicine agenda. *Nat Rev Genet* 22, 284–306. doi:
874 10.1038/s41576-020-00306-8
- 875 Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D.
876 (2006). Principal components analysis corrects for stratification in genome-wide
877 association studies. *Nature Genetics* 38, 904–909. doi: 10.1038/ng1847
- 878 Pruim, R. J., Welch, R. P., Sanna, S., Teslovich, T. M., Chines, P. S., Gliedt, T. P., et al. (2010).
879 LocusZoom: regional visualization of genome-wide association scan results.
880 *Bioinformatics* 26, 2336–2337. doi: 10.1093/bioinformatics/btq419

- 881 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al.
882 (2007). PLINK: a tool set for whole-genome association and population-based linkage
883 analyses. *American journal of human genetics* 81, 559–575. doi: 10.1086/519795
- 884 Ramsay, M., Crowther, N., Tambo, E., Agongo, G., Baloyi, V., Dikotope, S., et al. (2016).
885 H3Africa AWI-Gen Collaborative Centre: a resource to study the interplay between
886 genomic and environmental risk factors for cardiometabolic diseases in four sub-
887 Saharan African countries. *Global health, epidemiology and genomics* 1, e20–e20. doi:
888 10.1017/gh.2016.17
- 889 Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., et
890 al. (2015). An integrated map of structural variation in 2,504 human genomes. *Nature*
891 526, 75–81. doi: 10.1038/nature15394
- 892 Teumer, A., Li, Y., Ghasemi, S., Prins, B. P., Wuttke, M., Hermle, T., et al. (2019). Genome-
893 wide association meta-analyses and fine-mapping elucidate pathways influencing
894 albuminuria. *Nature Communications* 10, 4130. doi: 10.1038/s41467-019-11576-0
- 895 Teumer, A., Tin, A., Sorice, R., Gorski, M., Yeo, N. C., Chu, A. Y., et al. (2016). Genome-
896 wide Association Studies Identify Genetic Loci Associated With Albuminuria in
897 Diabetes. *Diabetes* 65, 803–817. doi: 10.2337/db15-1313
- 898 Tin, A., and Köttgen, A. (2020). Genome-Wide Association Studies of CKD and Related
899 Traits. *Clinical Journal of the American Society of Nephrology*, CJN.00020120. doi:
900 10.2215/CJN.00020120
- 901 Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic
902 variants from high-throughput sequencing data. *Nucleic Acids Research* 38, e164–
903 e164. doi: 10.1093/nar/gkq603
- 904 Wuttke, M., Li, Y., Li, M., Sieber, K. B., Feitosa, M. F., Gorski, M., et al. (2019). A catalog of
905 genetic loci associated with kidney function from analyses of a million individuals.
906 *Nature Genetics* 51, 957–972. doi: 10.1038/s41588-019-0407-x
- 907 Yang, J., Ferreira, T., Morris, A. P., Medland, S. E., Madden, P. A. F., Heath, A. C., et al.
908 (2012). Conditional and joint multiple-SNP analysis of GWAS summary statistics

- 909 identifies additional variants influencing complex traits. *Nature Genetics* 44, 369–375.
910 doi: 10.1038/ng.2213
- 911 Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: a tool for genome-
912 wide complex trait analysis. *American journal of human genetics* 88, 76–82. doi:
913 10.1016/j.ajhg.2010.11.011
- 914 Zanetti, D., Rao, A., Gustafsson, S., Assimes, T. L., Montgomery, S. B., and Ingelsson, E.
915 (2019). Identification of 22 novel loci associated with urinary biomarkers of albumin,
916 sodium, and potassium excretion. *Kidney International* 95, 1197–1208. doi:
917 10.1016/J.KINT.2018.12.017
- 918 Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for
919 association studies. *Nature genetics* 44, 821–824. doi: 10.1038/ng.2310
- 920 Zolotarov, Y., Ma, C., González-Recio, I., Hardy, S., Franken, G. A. C., Uetani, N., et al.
921 (2021). ARL15 modulates magnesium homeostasis through N-glycosylation of
922 CNNMs. *Cell Mol Life Sci* 78, 5427–5445. doi: 10.1007/s00018-021-03832-8
- 923
- 924
- 925

Studies – Genotyping - GWAS



Meta-Analysis Strategy



Post-GWAS Analysis and Polygenic Scores

